

Area law in the exact solution of many-body localized systems

Evgeny Mozgunov*

Institute for Quantum Information and Matter, Department of Physics, California Institute of Technology

(Dated: August 29, 2017)

Many-body localization was proven under realistic assumptions by constructing a quasi-local unitary rotation that diagonalizes the Hamiltonian (Imbrie, 2016). A natural generalization is to consider all unitaries that have a similar structure. We bound entanglement for states generated by such unitaries, thus providing an independent proof of area law in eigenstates of many-body localized systems. An error of approximating the unitary by a finite-depth local circuit is obtained. We connect the defined family of unitaries to other results about many-body localization (Kim et al, 2014), in particular Lieb-Robinson bound. Finally we argue that any Hamiltonian can be diagonalized by such a unitary, given it has a slow enough logarithmic lightcone in its Lieb-Robinson bound.

I. INTRODUCTION

Physical models change their behavior in the presence of disorder — a randomness in the terms of the corresponding Hamiltonian that's breaking the translation symmetry. First step in understanding the physics of disorder was made by Anderson³. He mathematically demonstrated a phenomenon of localization for non-interacting particles in a disordered potential. A quantum particle doesn't leave the vicinity of its starting point. This behavior of non-interacting particles can be thought of as a zeroth order approximation in the description of the interacting system. The question of convergence of any expansion around the non-interacting solution turned out to be a hard nut to crack. It was noted⁴ that an interacting system can be thought of as a single particle traveling in the Fock space, but the corresponding potential was not localizing in the Anderson sense. First positive answer was presented by Basko et al⁵, which is when this problem became widely known as many-body localization (MBL). The achievement of this pioneering work was to find a realistic system (electrons in grains with a weak tunneling between them) and a range of parameters (weak interaction) where the diagrammatic expansion becomes tractable. The authors demonstrated zero conductivity at a finite temperature — the system behaved as an Anderson insulator would.

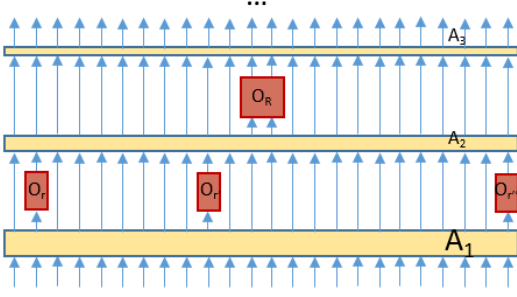
An alternative route of investigation concentrated around disordered spin systems and conjectures by Huse et al⁶. Spins-1/2 are described by Pauli matrices σ_i^α . These degrees of freedom are normally fixed, and excitations do not always admit a quasiparticle description; what localization meant was open to debate. Huse et al conjectured that dressed ("tailed" here and below) spin operators τ_i^α can be found that are conserved quantities of the dynamics. The set of those quasi-local conserved quantities τ_i^α was conjectured to form an alternative (logical) operator basis. The coefficients of the mapping between the two thus is given by the trace norm $\text{tr}(\sigma_i^\alpha \tau_j^\beta)$ which decays exponentially according to quasi-locality of τ_i^α . This mapping allows to rewrite the spin Hamiltonian as a classical Hamiltonian (only containing τ_i^z) in the logical basis. This logical Hamiltonian contains a non-local

exponentially decaying interaction inherited from quasi-locality of τ_i^α .

All of the above conjectures turn out to be true for a specific spin system solved by Imbrie¹. His construction directly obtained the logical Hamiltonian as expected by Huse et al⁶ by interleaving steps of perturbation theory with spatially sparse non-perturbative rotations. The exponential tails are adjusted as no decay happens in the regions containing those non-perturbative rotations, called resonances. But resonances are proven to be rare thus exponential decay holds on average. We will present the construction in more detail below, as it inspired our own definition of many-body localization. The path to showing quasi-locality of logical operators lies through a property called telescopic sum that we show. To our knowledge Imbrie's work¹ is the only first principles proof for spin systems that serves as a beacon guiding other mathematical work⁷ in the field.

Quite a lot of progress has been made by taking a different starting point than first principles (the Hamiltonian). Most notably, a Lieb-Robinson bound has been proven² from conjectures alike ones by Huse et al⁶. Lieb-Robinson bound characterizes how an operator evolves under the dynamics generated by the Hamiltonian. Normally the region on which the majority of the operator weight can be found (in terms of trace coefficients with the operator basis as shown above) grows linearly with evolution time. In MBL systems the operator spreads slower, only logarithmic in time. We show that is also true for the first principles solution.

The intuition about MBL was often informed by numerical simulations, in particular Nayak et al⁸ conjectured that a finite-depth local circuit can map an eigenstate approximately to a product state. We will confirm that conjecture by providing a bound on the error of such mapping, and emphasize that a single unitary works equally well for all eigenstates. A circuit on top of a product state can be thought of as a matrix-product state. Some of the authors focused on a so called strong MBL, where they were able to prove⁹ a matrix-product state representation of every eigenstate. They also conjectured that a stronger Lieb-Robinson bound where the lightcone does not spread at all holds for the states at

FIG. 1. Definition of the circuit U

low energies. We find no evidence for that claim.

II. PROPERTIES OF THE FAMILY OF IMBRIE'S SOLUTIONS

Every Hamiltonian H is characterized by its eigenvalues H_{diag} and eigenvectors U :

$$UHU^\dagger = H_{diag} \quad (1)$$

Imbrie¹ presented a construction of $U = \prod_k U_k$ that works for a specific Hamiltonian on a 1d chain:

$$H = \gamma \sum_i \sigma_i^x + \sum_i h_i \sigma_i^z + J_i \sigma_i^z \sigma_{i+1}^z \quad (2)$$

with $-1 \leq h_i, J_i \leq 1$ uniform i.i.d. random numbers and $\gamma \ll 1$. The structure of $U = \prod_k U_k$ is illustrated in Figure 1. We shall present the construction abstractly, without specifying all of the internal structure of U_k . We believe that for any MBL system a generalization of Imbrie's construction can be defined, so our presentation can be read both as a restatement of Imbrie's proof, and as a postulate about any system sufficiently deep in the MBL phase.

First consider no resonances (no red blocks in Figure 1). There are infinitely many terms in the product $U = \prod_k U_k$ counted by $k = 1, 2, \dots$

Postulate II.1. (Support) Each block $U_k = e^{A^{(k)}}$, where $iA^{(k)}$ is a Hermitean generator that is a sum of $\text{floor}(\frac{15}{7}L_k)$ -local terms on each site of the chain, where $L_k = (15/8)^k$:

$$A^{(k)} = \sum_i A_i^{(k)} \quad (3)$$

Here locality is in terms of the interval on the chain that operator covers — the distance in space between the leftmost and the rightmost spin that the operator depends on. We provide a guide to the text of Imbrie's paper in the Appendix B, explaining how we chose the number $15/7$ in the support definition above. We also conclude that Imbrie's proof allows one to restrict the norm of $\|A_i^{(k)}\|$ as follows:

Postulate II.2. (Norm)

$$\|A_i^{(k)}\| \leq c' \chi^{\text{ceil} L_{k-1}} \quad (4)$$

where $\chi = c\gamma^{0.95}$ and c, c' are constant numbers.

Note that the magnetic field γ enters the definition here. The construction rests on the fact that the $\frac{15}{7}L_k$ appearing in the support has the same scaling with k as $L_{k-1} = \frac{8}{15}L_k$. The coefficient c is obtained as a polynomial of $\exp(\text{supp} A_i^{(k)}/L_k)$. If scalings were to be different even by a logarithmic factor (say, supports grow as kL_k), then instead of $(c\gamma^{0.95})^{L_{k-1}}$ the bound would contain $L_{k-1}! \gamma^{L_{k-1}}$, which diverges regardless how small γ is. In Imbrie's paper the fact that support has $\sim L_k$ scaling is used several times, however the proof of it raises some questions, as we note in Appendix B. These two postulates are enough to prove the convergence of the no-resonance infinite product $U = \prod_k U_k$ in the following sense:

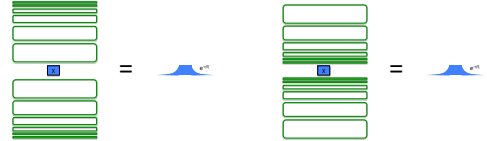
Theorem II.1 (collar approximation). *Operators acquire exponentially decaying tails when conjugated by no-resonance U .*

$$\|U XU^\dagger - U_c XU_c^\dagger\| \leq 24 \cdot 2^{0.5S(X)} \alpha^{c+1} \|X\|, \quad (5)$$

$$\|U^\dagger XU - U_c^\dagger XU_c\| \leq 24 \cdot 2^{0.5S(X)} \alpha^{c+1} \|X\| \quad (6)$$

where U_c is constructed using the generators $A^{(k)}$ truncated to a collar of c sites of the support of X (all the terms reaching outside the collar are dropped). $\alpha = 24c' \chi^{\frac{56}{225}}$ is a power of transverse field γ .

For the proof, see Appendix A.



Consider the implications of this result for the eigenvalues H_{diag} .

$$H_{diag} = UHU^\dagger = \gamma \sum_i (\sigma_i^x)_{tailed} + \sum_i h_i (\sigma_i^z)_{tailed} + J_i (\sigma_i^z \sigma_{i+1}^z)_{tailed}$$

On the other hand, since H_{diag} is a diagonal matrix:

$$H_{diag} = f(\sigma_z) = \sum_{s \subset \{1 \dots L\}} J_s \prod_{i \in s} \sigma_i^z. \quad (7)$$

The bound on J_s can be directly inferred from the general bound on the telescopic sum (Theorem A.2 in the Appendix):

Corollary II.1.1.

$$J_s \leq 12\alpha^{d(s)} \quad (8)$$

where $\alpha = 24c' \chi^{\frac{56}{225}}$, $d(s)$ is consecutive support and no resonances are within it.

Using a time-averaged version of this (see Section III) as a starting point in an ensemble-averaged sense, Kim et al² derive Lieb-Robinson bound:

Corollary II.1.2.

$$\begin{aligned} \|[A, B(t)]\| &\leq ate^{-bd_{A,B}} = \\ &= at\alpha^{d_{A,B}} \end{aligned}$$

where $B(t) = e^{iHt}Be^{-iHt}$, $\alpha = 24c'\chi^{\frac{56}{225}}$, $a = c\|A\|\|B\|$ and c is a constant number.

We provide an independent derivation in Appendix A. Another known MBL effect of logarithmic growth of entanglement across the cut can also be proven². We prove that U is close to a finite-depth local circuit in Appendix A.2.

The benefit of taking Postulate II.2 as a starting point is that one gets area law essentially for free. Consider an eigenstate $|E_s\rangle = U^\dagger|s\rangle$ where $|s\rangle$ is a product state of spins in z-basis and a cut at some location:

Theorem II.2 (Area law).

$$S_{cut} = S(\rho_{left}) < 650c'\chi$$

where $\rho_{left} = \text{tr}_{right}|E_s\rangle\langle E_s|$.

No other proofs of area law in an eigenstate appeared in the literature to our knowledge. It was considered to be an obvious corollary of exponential decay of correlations, however Imbrie did not prove exactly the same decay of correlations as required for 1d area law.¹⁰ We show our proof in Appendix A, and obtain the Postulate II.2 from the construction by Imbrie¹ in Appendix B.

We now interleave the circuit with *resonances* — locations where perturbation theory in γ failed and a local non-perturbative rotation had to be applied.

Postulate II.3. (*R-size*) At each step, the operators acting between $e^{A^{(k)}}$ are

$$O_k = \prod_r O_{k,r}, \quad (9)$$

where $O_{k,r}$ are big unitary rotations supported on $< 4.2L_k = 4.2(15/8)^k$ consecutive spins.

Let R_k be the set of spins involved in O_k . Imbrie¹ explicitly presents a following bound on how rare they are:

Postulate II.4. (*R-density*) The density of step- k resonances

$$P(i \in R_k) = \lim_{L \rightarrow \infty} \frac{|R_k|}{L} \leq c\epsilon^{c' \log^2 L_k} \quad (10)$$

where $\epsilon = \gamma^{0.05}$ and c, c' are constant numbers.

We postulate 1-4 as a definition of a full U . Now entanglement bound strongly depends on the position of the cut!

Corollary II.2.1.

$$\overline{S_{cut}} \leq 650c'\chi + 4\epsilon \ln 2$$

and $\max(S_{cut}) \sim \log L$

Here $\overline{S_{cut}}$ is the average over different positions in the system (equivalently, different realizations of disorder), and the true finite size scaling is more complicated than $\log L$ but is slower than any power. We note an incredibly small power in the resonance density bound $\epsilon = \gamma^{0.05}$. That means that Imbrie's construction as is cannot be applied for magnetic fields $\gamma \geq 10^{-6}$. MBL survives for much bigger magnetic fields, and this discrepancy is due to multiple overestimations needed for Imbrie's proof to got through. If one uses a more strict choice of which regions to call resonant, their density remains controlled roughly up to the transition to extended phase. We demonstrate it numerically in a companion paper^{11,12}.

III. TIME-AVERAGING

Consider the following transformation of a local operator X :

$$X' = \lim_{T \rightarrow \infty} 1/T \int_0^T e^{iHt} X e^{-iHt} dt; \quad (11)$$

this kills off-diagonal elements in the eigenbasis of H , so commutation with H is enforced. Time-averaging is used in Kim et al² as part of a definition of MBL. They show the following for time-averaging with an MBL Hamiltonian:

Lemma III.1. *Infinite time-averaging introduces exponential tails.*

$$\|X' - X'_c\| \leq a(4\alpha)^c \quad (12)$$

where X'_c is the truncation of X' to a region with collar c around the support of X , $\alpha = 24c'\chi^{\frac{56}{225}}$ and a, Z are constant numbers.

Proof. Following², logical spin operators $U\sigma_i^\alpha U^\dagger$ form an operator basis and obey the telescopic sum bound proven in Appendix A. Coefficients of an operator X in this basis are given by traces of products $\text{tr} X U\sigma_i^\alpha U^\dagger$, which are bounded as $c'\alpha^{i-x}$ where i_X is the nearest site in X . Only the coefficients of combinations of $U\sigma_i^\alpha U^\dagger$ survive the infinite time averaging. Then each term has a telescopic bound $c'\alpha^c$ where c is the size of the collar. Two bounds combine to give a collar bound on X' . The α is only increased by a square of local dimension as there are many possible operator strings. \square

If we time-average a local patch of the Hamiltonian H_a in $H = H_a + H_b$ for a finite-time, the commutation with

H is not fully enforced. Let's bound the commutator of the two parts:

$$[H'_a, H'_b] = [H'_a, H] = \frac{1}{T} \int_0^T e^{iHt} [H_a, H] e^{-iHt} dt = \quad (13)$$

$$= \frac{1}{iT} \int_0^T \frac{d}{dt} e^{iHt} H_a e^{-iHt} dt = \frac{e^{iHT} H_a e^{-iHT} - H_a}{iT} \quad (14)$$

$$\|[H'_a, H'_b]\| \leq \frac{2S}{T}, \quad (15)$$

where S is the support of H_a . The spread of the local operator of any Hamiltonian can be bounded by the Lieb-Robinson bound. Let the light cone (where the bound on $\|[A(t), B]\| \sim 1$) be given by $x(t)$, therefore time-averaging for time T should give an operator whose support outside the radius $x(T) + \Delta x$ can be bounded as $e^{-c\Delta x}$, where c is a Hamiltonian-independent number for 2-local Hamiltonians (and can be strengthened even more for MBL Lieb-Robinson bound: $e^{-c\Delta x} = \alpha^{\Delta x}$ now H -dependant unlike the generic Lieb-Robinson bound). Moreover, the time-averaging can be done with the Hamiltonian truncated to an $x(T) + \Delta x$ collar of the local operator, with only exponentially small error:

$$\tilde{H}'_a = 1/T \int_0^T e^{iH_c t} H_a e^{-iH_c t} dt = H'_a + \delta_a, \quad \|\delta_a\| \leq e^{-c\Delta x}. \quad (16)$$

Note that δ is a nonlocal interaction that's exponentially small in the distance as e^{-cd} . Now split the system in many regions $a_1, a_2 \dots$. So if we take H_a to have support $S > x(T) + \Delta x$, the effects of the time-averaging could be truncated in such a way that \tilde{H}'_{a_i} only overlap with their neighbors $\tilde{H}'_{a_{i\pm 1}}$. So we have reduced our Hamiltonian to a nearest neighbor chain of approximately commuting terms (as in Bound 15), with an exponentially small error:

$$H = \sum_i \tilde{H}'_{a_i} + \delta_i. \quad (17)$$

Now let's take even-odd pairs $\tilde{H}'_{a_{2k}} + \tilde{H}'_{a_{2k+1}}$. We would like to use a result by¹³ that reduces approximately commuting operators to exactly commuting ones, by small

adjustments to each:

$$\tilde{H}'_{a_{2k}} = A_{2k} + \epsilon_{2k}, \quad \tilde{H}'_{a_{2k+1}} = A_{2k+1} + \epsilon_{2k+1}, \quad (18)$$

$$[A_{2k}, A_{2k+1}] = 0, \quad (19)$$

and the $\|\epsilon\| < (2S/T)^{1/6}$ and have the same support as the overlap of a'_{2k} and a'_{2k+1} (the supports of corresponding H'). Now we introduce the notation

$$h_k = A_{2k} + A_{2k-1}, \quad [h_k, h_{k+1}] = 0, \quad (20)$$

$$H = \sum_k h_k + \sum_i (\epsilon_i + \delta_i). \quad (21)$$

We have reduced any local Hamiltonian on a chain to a nonlocal one with exponentially decaying long-range interaction, written in the form of commuting terms plus a small perturbation. We note that to repeat this in 2d one would need a version of Hastings' result for three matrices, and that one has a counterexample presented in his paper.

For chains, one may think of applying Imbrie's construction to this form of the Hamiltonian - after step 1, Imbrie's Hamiltonian has exactly this form with long range but rapidly decaying interactions! Unfortunately, the bounds for a general 1d Hamiltonian do not work out in our favor: the number of levels in the block $S > x(T) + \Delta x$ grows as 2^S , so the minimum level distances decrease as 2^{-S} . However, the norm bound on ϵ is just a power $1/T$. For the standard Lieb-Robinson bound $x(T) = vT$ and $S > x(T)$, so the level distances are always smaller than the perturbation scale, which is the opposite to what is used in Imbrie's construction. However, in MBL with $x(T) \sim \ln T$, the scales may just work out: $T \sim (1/\alpha)^{x(T)}$, so the norm of the perturbation goes as $\alpha^{\tilde{c}S}$ where S is our choice of the support and \tilde{c} is some constant. It decreases faster than $(1/2)^S$ for a sufficiently small α . We have shown the following:

Corollary III.1.1. *If H obeys MBL Lieb-Robinson bound*

$$\|[A, B(t)]\| \leq at\alpha^{d_{A,B}}, \quad (22)$$

then the perturbation on top of a S -local commuting component of H $V = \sum_i V_i$ is sufficiently small for Imbrie's proof to go through:

$$\|V_i\| \leq c'' \alpha^{c'S} \quad (23)$$

where c'', c', S are constant numbers.

In this way, any Hamiltonian that possesses the MBL Lieb-Robinson bound can be reduced to a form "diagonal + small perturbation", even if the z -basis is not apparent from the start.

* mvjenia@gmail.com

¹ J. Z. Imbrie, "On many-body localization for quantum spin

- chains,” J. Stat. Phys. **163**, 998 (2016).
- ² I. H. Kim, A. Chandran, and D. A. Abanin, “Local integrals of motion and the logarithmic lightcone in many-body localized systems,” e-print arXiv:1412.3073 (2014).
 - ³ P. W. Anderson, “Absence of diffusion in certain random lattices,” Phys. Rev. **109**, 1492 (1958).
 - ⁴ D. J. Thouless, “Maximum metallic resistance in thin wires,” Physical Review Letters **39**, 1167–1169 (1977).
 - ⁵ D. Basko, I. Aleiner, and B. Altshuler, “Metalinsulator transition in a weakly interacting many-electron system with localized single-particle states,” Annals of Physics **321**, 11261205 (2006).
 - ⁶ D. A. Huse, R. Nandkishore, and V. Oganesyan, “Phenomenology of fully many-body-localized systems,” Phys. Rev. B **90** (2014).
 - ⁷ D. A. Abanin, W. D. Roeck, and F. Huveneers, “Theory of many-body localization in periodically driven systems,” Annals of Physics **372**, 1–11 (2016).
 - ⁸ B. Bauer and C. Nayak, “Area laws in a many-body localized state and its implications for topological order,” J. Stat. Mech. , P09005 (2013).
 - ⁹ M. Friesdorf, A. Werner, W. Brown, V. Scholz, and J. Eisert, “Many-body localization implies that eigenvectors are matrix-product states,” Physical Review Letters **114**, 170505 (2015).
 - ¹⁰ F. G. S. L. Brandão and M. Horodecki, “An area law for entanglement from exponential decay of correlations,” Nature Physics **9**, 721–726 (2013).
 - ¹¹ E. Mozgunov, “Avoided crossings are resonances in a many-body localized system,” to be published (2017).
 - ¹² E. Mozgunov, *Slow Drive of Many-Body Localized Systems*, PhD thesis, Caltech, Department of Physics (2017).
 - ¹³ M. B. Hastings, “Making almost commuting matrices commute,” Communications in Mathematical Physics **291**, 321345 (2009).
 - ¹⁴ A. M. Childs, D. W. Leung, and G. Vidal, “Reversible simulation of bipartite product hamiltonians,” IEEE Transactions on Information Theory **50** (2004).
 - ¹⁵ S. Bravyi, M. B. Hastings, and F. Verstraete, “Lieb-robinson bounds and the generation of correlations and topological quantum order,” Physical Review Letters **97** (2006).
 - ¹⁶ J. Eisert and T. J. Osborne, “General entanglement scaling laws from time evolution,” Physical Review Letters **97**, 150404 (2006).
 - ¹⁷ A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and quantum computation*, Graduate Studies in Mathematics, Vol. 47 (American Mathematical Society, 2002).

Appendix A: Proofs of properties of Imbrie Circuit

1. Telescopic sum for Imbrie circuit

a. Small unitary expansion

We consider a generator $A^{(k)}$ of a unitary evolution $e^{A^{(k)}}$ (corresponding to step k) such that $A^{(k)} = \sum_i A_i^{(k)}$ where $\|A_i^{(k)}\| \leq c' \chi^{\text{ceil} L_{k-1}}$ and each A_i is supported on $\text{floor}(\frac{8}{7} L_{k+1})$ consecutive sites. We would like to study the structure of a local operator X , evolved as $e^{A^{(k)}} X e^{-A^{(k)}}$. For the purposes of the proofs below, we use the term support $S(X)$ of any operator X as the shortest interval of the chain containing all the sites where action of X is distinct from identity. We prove

Lemma A.1.

$$e^{A^{(k)}} X e^{-A^{(k)}} = \sum_{j=0}^{\infty} X_j \quad (\text{A1})$$

where $S(X_j) \leq S(X) + j\Delta S$, $\Delta S = 2(\text{floor}(\frac{8}{7} L_{k+1}) - 1)$ and

$$\|X_j\| \leq 2^{j_0} (2c' \Delta S)^j \chi^{j \text{ceil} L_{k-1}} \|X\| \quad (\text{A2})$$

where $j_0 = \text{ceil} \frac{S(X) + S(A) - 2}{\Delta S}$ and $S(A) = \text{floor}(\frac{8}{7} L_{k+1})$

So the operator X acquires exponential tails after conjugating with $e^{A^{(k)}}$.

Proof. We drop the superscript (k) in $A^{(k)}$ for convenience.

$$e^A X e^{-A} = \sum_n \frac{A^n}{n!} X \sum_m \frac{(-A)^m}{m!} \quad (\text{A3})$$

collecting the terms that have the same power of A

$$= \sum_j \sum_{m=0}^j \frac{A^{j-m}}{(j-m)!} X \frac{(-A)^m}{m!} \quad (\text{A4})$$

Compare this to the expansion of j 'th order commutator

$$[A[A \dots [A, X]]]^{(j)} = \sum_{m=0}^j C_m^j A^{j-m} X (-A)^m \quad (\text{A5})$$

where $C_m^j = \frac{j!}{(j-m)!m!}$

$$\sum_j \sum_{m=0}^j \frac{A^{j-m}}{(j-m)!} X \frac{(-A)^m}{m!} = \sum_j \frac{1}{j!} [A[A \dots [A, X]]]^{(j)} \quad (\text{A6})$$

Now denote $X_j = \frac{1}{j!} [A[A \dots [A, X]]]^{(j)}$. We found that

$$e^A X e^{-A} = \sum_j X_j \quad (\text{A7})$$

Recall that support of X is $S(X)$ consecutive sites and the support of A is $S(A) = \text{floor} \frac{8}{7} L_{k+1}$. A commutator $[A, O]$ for a local operator O of support $S(O)$ consecutive sites is nonzero for only the terms in $A = \sum_i A_i$ that overlap with O . The index i goes over every site, so there are $S(O) + S(A) - 2$ such terms. The support of $[A, O]$ becomes $S(O) + 2S(A) - 2$. The support of X_j which is the j 'th order of the commutator just grows linearly:

$$S(X_j) = S(X) + j(2S(A) - 2) = S(X) + j\Delta S \quad (\text{A8})$$

where we denoted $\Delta S = 2S(A) - 2$. The number of terms of A that have nonzero commutator at j 'th step is given by $S(O) + S(A) - 2$ for $O = X_{j-1}$:

$$n_j = S(X_{j-1}) + S(A) - 2 = S(X) + j\Delta S + S(A) - 2 \quad (\text{A9})$$

For the norm bound on X_j , a prefactor before the j 'th power of $\|A\|$ is $1/j!$ and a product of all n_j 's:

$$\|X_j\| \leq \frac{1}{j!} \left(\prod_{i=1}^j n_i \right) \|A\|^j \|X\| \quad (\text{A10})$$

Let us bound the prefactor:

$$\frac{1}{j!} \prod_{i=1}^j n_i = \frac{1}{j!} \prod_{i=1}^j (S(X) + S(A) - 2 + i\Delta S) = \frac{\Delta S^j}{j!} \prod_{i=1}^j \left(\frac{S(X) + S(A) - 2}{\Delta S} + i \right) \leq \quad (\text{A11})$$

$$\leq \frac{\Delta S^j}{j!} \prod_{i=1}^j \left(\text{ceil} \frac{S(X) + S(A) - 2}{\Delta S} + i \right) = \frac{\Delta S^j}{j!} \prod_{i=1}^j (j_0 + i) = \Delta S^j \frac{(j + j_0)!}{j! j_0!} = \Delta S^j C_{j_0}^{j+j_0} \quad (\text{A12})$$

where we denoted $j_0 = \text{ceil} \frac{S(X) + S(A) - 2}{\Delta S}$. Finally we use $C_{j_0}^{j+j_0} \leq 2^{j+j_0}$ to arrive at

$$\frac{1}{j!} \prod_{i=1}^j n_i \leq (2\Delta S)^j 2^{j_0} \quad (\text{A13})$$

Combining that with the bound on $\|A\|$, we get the desired form:

$$\|X_j\| \leq 2^{j_0} (2c'\Delta S)^j \chi^{j \text{ceil} L_{k-1}} \|X\| \quad (\text{A14})$$

□

a. Aside: comparison with Lieb-Robinson Bound We have found that if one conjugates a local operator X with $S(A)$ -locally generated 1d evolution e^A for small time $t \leq c'\chi^{\text{ceil} L_{k-1}}$, operator X acquires exponential tails.

It is a much stronger result than if we apply Lieb-Robinson Bound for small χ for propagation of tails beyond the collar, but becomes trivial for $\chi = 1$. Here's this LRB estimate for comparison - it still gives nontrivial answers even for big χ :

$$\sum_{j=c}^{\infty} \|X_j\| \leq \exp\{-(ac - S(A)c'\chi^{\text{ceil} L_{k-1}})\} \|X\| \quad (\text{A15})$$

Here a is just some number < 20 much smaller than $\text{ceil} L_{k-1}/(-\ln \chi)$. We shall use the stronger first bound as in our case $\chi \ll 1$.

b. Infinite product of unitaries

Now we'd like to use our bound to estimate precision of collar approximations to:

$$\prod_{k=1}^{\infty} e^{A^{(k)}} X \prod_{k'=1}^{\infty} e^{-A^{(k')}} \quad (\text{A16})$$

as in the total Imbrie circuit without resonances. The order in the product will turn out to be unimportant for our bound. From the point of view of the physical meaning, if smaller k act on the operator X first, then X is a physical operator and evolved X represents the action on "logical" spins, or integrals of motion. In particular, for the diagonalization of the Hamiltonian, smaller k act on the physical Hamiltonian first. If smaller k act last, then X is a logical operator, and its evolved form will be the representation of such operator on physical spins.

We now present and prove the bound:

Theorem A.2.

$$\prod_{k=1}^{\infty} e^{A^{(k)}} X \prod_{k'=1}^{\infty} e^{-A^{(k')}} = \sum_{j=0}^{\infty} X_j, \quad S(X_j) = S(X) + 2j, \quad \|Y_n\| \leq 6 \cdot 2^{0.5S(X)} \alpha^j \|X\| \quad (\text{A17})$$

where $\alpha = 24c' \chi^{\frac{56}{225}}$ and the order in the product is unimportant.

If we add a collar of c sites to the support of X , and take $A_c^{(k)}$ to be the sum of terms fully within the collared region, then

$$\left\| \prod_{k=1}^{\infty} e^{A^{(k)}} X \prod_{k'=1}^{\infty} e^{-A^{(k')}} - \prod_{k=1}^{\infty} e^{A_c^{(k)}} X \prod_{k'=1}^{\infty} e^{-A_c^{(k')}} \right\| \leq 24 \cdot 2^{0.5S(X)} \alpha^{c+1} \|X\| \quad (\text{A18})$$

$\alpha \leq 0.5$ is required for the second bound to hold.

Proof. We start from looking at individual rotation, that was discussed above:

$$U_k X U_k^\dagger = \sum_{j_k} X_{j_k}, \quad S(X_{j_k}) = S(X) + 2j_k \text{floor}\left(\frac{8}{7}L_{k+1} - 1\right), \quad \|X_{j_k}\| \leq 2^{j_0} (2c' \Delta S)^j \chi^{j \text{ceil} L_{k-1}} \|X\| \quad (\text{A19})$$

If there are two rotations, we get:

$$U_{k'} U_k X U_k^\dagger U_{k'}^\dagger = \sum_{j_k, j'_k} X_{j_k, j'_k}, \quad (\text{A20})$$

$$S(X_{j_k, j'_k}) = 2j_k \text{floor}\left(\frac{8}{7}L_{k+1} - 1\right) + 2j'_k \text{floor}\left(\frac{8}{7}L_{k'+1} - 1\right), \quad (\text{A21})$$

$$\|X_{j_k, j'_k}\| \leq 2^{j_0, k + j_0, k' (j_k)} (2c' \Delta S_k)^{j_k} (2c' \Delta S'_k)^{j'_k} \chi^{j_k \text{ceil} L_{k-1} + j'_k \text{ceil} L_{k'-1}} \|X\| \quad (\text{A22})$$

we see that the order of unitaries indeed does not matter except for the first term, that depends on the support of previous terms. We will see in paragraph A 1 b a how that term will remain bounded for every X_{j_k, j'_k} s.t. $S(X_{j_k, j'_k}) = S(X) + 2j$ regardless of the order. For an infinite product, the components of evolved X are indexed by a string of integers $\{j_k\}$ for all $k = 1, 2, 3 \dots$:

$$\prod_{k=1}^{\infty} e^{A_k} X \prod_{k'=1}^{\infty} e^{-A_{k'}} = \sum_{\{j_k\}} X_{\{j_k\}}, \quad (\text{A23})$$

$$S(X_{\{j_k\}}) = S(X) + \sum_k 2j_k \text{floor}\left(\frac{8}{7}L_{k+1} - 1\right), \quad (\text{A24})$$

$$\|X_{\{j_k\}}\| \leq C(\{j_k\}, \text{order}) \left(\prod_k (2c' \Delta S_k)^{j_k} \right) \chi^{\sum_k j_k \text{ceil} L_{k-1}} \|X\| \quad (\text{A25})$$

$$C(\{j_k\}, \text{order}) = 2^{j_0, k(1) + j_0, k(2) (j_{k(1)}) + \dots} \quad (\text{A26})$$

Terms $X_{\{j_k\}}$ whose resulting supports $S(X_{\{j_k\}}) = S(X) + 2j$ ($j \in \mathbb{Z}$) are combined into X_j :

$$X_j = \sum_{\{j_k\} | S(X_{\{j_k\}}) = S(X) + 2j} X_{\{j_k\}} \quad (\text{A27})$$

We will provide a uniform bound in terms of j on the norm of each $X_{\{j_k\}}$ in X_j , and also argue that the total number of terms is only exponential in j .

One of the terms contributing to $\|X_{\{m_k\}}\|_j$ is

$$A = \prod_k (2c' \Delta S_k)^{j_k} \quad (\text{A28})$$

We know that $\Delta S_k = 2\text{floor}(\frac{8}{7}L_{k+1} - 1)$. We will leave one of the powers of two to the $C(\{j_k\}, \text{order})$, and bound a

$$A = \prod_k \left(2c' \text{floor} \left(\frac{8}{7}L_{k+1} - 1 \right) \right)^{j_k} \leq \prod_k \left(\frac{16}{7}c' L_{k+1} \right)^{j_k} \quad (\text{A29})$$

The collar is increased in steps of $S(A_k) - 1 = \text{floor} \frac{8}{7}L_{k+1} - 1$ on each side, so

$$j = \sum_k (S(A_k) - 1)j_k \quad (\text{A30})$$

We can collect two sums over j_k in our bound:

$$A \leq \left(\frac{16}{7}c' \right)^{\sum_k j_k} \left(\frac{15}{8} \right)^{\sum_k (k+1)j_k} \quad (\text{A31})$$

Note that

$$\sum_k j_k \leq \sum_k (k+1)j_k \leq \sum_k (S(A_k) - 1)j_k = j \quad (\text{A32})$$

so assuming $c' > 1$ (or else it can be replaced by 1 in the original bound):

$$A \leq \left(\frac{30}{7}c' \right)^j \quad (\text{A33})$$

That's a good enough bound for our purposes. Now let's turn our attention to powers of χ :

$$\chi^{\sum_k j_k \text{ceil} L_{k-1}} \quad (\text{A34})$$

We'd like to bound it by a j -dependent expression as well. $\chi < 1$, so we need to bound the sums in the opposite direction:

$$\sum_k \text{ceil} L_{k-1} j_k \geq \sum_k L_{k-1} j_k = \left(\frac{8}{15} \right)^2 \frac{7}{8} \sum_k \frac{8}{7} L_{k+1} j_k \geq \quad (\text{A35})$$

$$\geq \left(\frac{8}{15} \right)^2 \frac{7}{8} \sum_k \text{floor} \left(\frac{8}{7} L_{k+1} - 1 \right) j_k = \left(\frac{8}{15} \right)^2 \frac{7}{8} j \quad (\text{A36})$$

So the resulting power of χ is at least

$$\chi^{\sum_k j_k \text{ceil} L_{k-1}} \leq \chi^{\left(\frac{8}{15} \right)^2 \frac{7}{8} j} \quad (\text{A37})$$

Finally, we deal with the order-dependent factor:

a. 2^{j+j_0} factor The factor 2^{j_0} might be problematic for the proofs involving the full circuit, as X is oftentimes a big object, and the earlier steps expand it even bigger.

For the proof above, we needed the telescopic sum bound:

$$\rightarrow \|X_{j_k}\| \leq 2^{j_k+j_{0,k}} (c' \Delta S)^{j_k} \chi^{j_k \text{ceil} L_{k-1}} \|X\| \quad (\text{A38})$$

For convenience, we pulled out 2^j . All factors except the first are independent on the multiplication order and their contributions were bounded above. We will show that the product of $2^{j_k+j_{0,k}}$ over all steps is bounded by M^j for some constant M . The bound is not multiplication order-dependent.

The contribution of terms $2^{j_k+j_{0,k}}$ to a single representative of the X_j is:

$$\prod_{k=1}^{k_{max}} 2^{j_k+j_{0,k}} = 2^{\sum_{k=1}^{k_{max}} j_k+j_{0,k}} \quad (\text{A39})$$

Now we need to bound the sum:

$$\sum_{k=1}^{k_{max}} j_k + j_{0,k} = \sum_{k=1}^{k_{max}} j_k + \text{ceil} \frac{S(X_{k-1}) + S(A_k) - 2}{2S(A_k) - 2} \quad (\text{A40})$$

where we have introduced X_{k-1} as the term that we took from the previous step. The sequence of supports $S(X) = S(X_0), S(X_1) \dots S(X_{k_{max}}) = 2j + S(X_0)$ should be monotonic. $S(A_k) = \text{floor} \frac{8}{7} L_{k+1}$. Finally, j_k is the number of terms in the commutator we selected at step k , so $S(X_k) = S(X_{k-1}) + j_k(2S(A_k) - 2)$, so $j_k = (S(X_k) - S(X_{k-1})) / (2S(A_k) - 2)$

$$\sum_{k=1}^{k_{max}} j_k + j_{0,k} \leq \sum_{k=1}^{k_{max}} \frac{S(X_k) + S(A_k) - 2}{2S(A_k) - 2} + 1 \leq \frac{3}{2} k_{max} + (S(X_{k_{max}}) - 1) \sum_{k=1}^{k_{max}} \frac{1}{2S(A_k) - 2} \quad (\text{A41})$$

The series at the end is bounded:

$$\sum_{k=1}^{k_{max}} \frac{1}{2S(A_k) - 2} = \frac{1}{2} \sum_{k=1}^{k_{max}} \frac{1}{\text{floor} \frac{8}{7} L_{k+1} - 1} \leq \frac{1}{2} \sum_{k=1}^{\infty} \frac{1}{\frac{8}{7} L_{k+1} - 2} \leq \frac{1}{2} \sum_{k=1}^{\infty} \frac{1}{\frac{8}{7} L_k} = \frac{7}{30} \frac{1}{1 - \frac{8}{15}} = \frac{1}{2} \quad (\text{A42})$$

So we arrive at a bound:

$$\sum_{k=1}^{k_{max}} j_k + j_{0,k} \leq j + \frac{3k_{max} + S(X_0) - 1}{2} \quad (\text{A43})$$

we note that $k_{max} \sim \ln j$, so we indeed only get an extra power of j , which will be suppressed by a small enough χ . Let's recall the specific dependence of k_{max} on j . $X_0 = X$, then X_1 is everything within a one site collar, X_j - within a j -site long collar. The collar is increased in steps of $S(A_k) - 1 = \text{floor} \frac{8}{7} L_{k+1} - 1$ on each side, so $\text{floor} \frac{8}{7} L_{k_{max}+1} - 1 \leq j$:

$$k_{max} = \max_k \text{ s.t. } \text{floor} \frac{8}{7} L_{k+1} - 1 \leq j \quad (\text{A44})$$

$$\frac{8}{7} L_{k_{max}+1} - 2 \leq j, \quad (15/8)^{k_{max}} = L_{k_{max}} \leq \frac{7}{15} (j+2) \quad (\text{A45})$$

$$k_{max} \ln(15/8) \leq \ln(j+2) + \ln(7/15) \quad k_{max} \leq \frac{\ln((7/15)(j+2))}{\ln(15/8)} \leq \frac{(7/15)(j+2)}{\ln(15/8)} \quad (\text{A46})$$

since $\frac{7}{15}(j+2) > 1$. So our previous bound becomes:

$$\sum_{k=1}^{k_{max}} j_k + j_{0,k} \leq \frac{\left(2 + \frac{7}{5\ln(15/8)}\right) j + \frac{14}{5\ln(15/8)} - 1 + S(X_0)}{2} \leq 2.2j + 1.8 + 0.5S(X_0) \quad (\text{A47})$$

And the prefactor is bounded as:

$$\prod_{k=1}^{k_{max}} 2^{j_k+j_{0,k}} \leq 2^{2.2j+1.8+0.5S(X_0)} = (2^{2.2})^j 2^{1.8+0.5S(X_0)} \quad (\text{A48})$$

Now putting it all together, we've learned that

$$\|X_{\{j_k\}}\| \leq (2^{2.2})^j 2^{1.8+0.5S(X_0)} \left(\frac{30}{7} c'\right)^j \chi^{\left(\frac{8}{15}\right)^2 \frac{7}{8} j} \|X\| \quad (\text{A49})$$

for $X_{\{j_k\}} \in X_j$. To bound $X_j = \sum_{\{j_k\}|j} X_j$ we need to upper bound the number of terms in the sum.

b. Number of terms We need to estimate the number of terms with collar j that are generated in our telescopic sums $\sum_{\{j_k\}} X_{\{j_k\}}$. First note that terms with $j_k \neq 0$ for at least one $k > k_{max}$ do not contribute. We're bounding the number of ways

$$\sum_k j_k \left(\text{floor} \frac{8}{7} L_{k+1} - 1 \right) = j \quad (\text{A50})$$

can be achieved. Note that for each j_k

$$j_k \left(\text{floor} \frac{8}{7} L_{k+1} - 1 \right) \leq j \quad (\text{A51})$$

So there are at most $j / (\text{floor} \frac{8}{7} L_{k+1} - 1)$ choices for each j_k up to k_{max} . This gives an upper bound on the number of terms:

$$\sum_{\{j_k\}|j} 1 \leq \prod_{k=1}^{k_{max}} \frac{j}{\text{floor} \frac{8}{7} L_{k+1} - 1} = \frac{j^{k_{max}}}{\prod_{k=1}^{k_{max}} \text{floor} \frac{8}{7} L_{k+1} - 1} \quad (\text{A52})$$

For the upper bound on $\sum_{\{j_k\}|j} 1$, we need to lower bound the denominator:

$$\prod_{k=1}^{k_{max}} \text{floor} \frac{8}{7} L_{k+1} - 1 \geq \prod_{k=1}^{k_{max}} L_k = \left(\frac{15}{8} \right)^{\sum_{k=1}^{k_{max}} k} = \left(\frac{15}{8} \right)^{\frac{k_{max}(k_{max}+1)}{2}} \quad (\text{A53})$$

We obtain the bound on the number of terms as:

$$\sum_{\{j_k\}|j} 1 \leq j^{k_{max}} \left(\frac{8}{15} \right)^{\frac{k_{max}(k_{max}+1)}{2}} \quad (\text{A54})$$

Note that we will need the opposite inequality on k_{max} :

$$k_{max} = \max_k \text{ s.t. } \text{floor} \frac{8}{7} L_{k+1} - 1 \leq j \quad (\text{A55})$$

$$\text{floor} \frac{8}{7} L_{k_{max}+2} - 1 > j \quad (\text{A56})$$

$$L_{k_{max}+2} > \frac{7}{8} j \quad (\text{A57})$$

$$\left(\frac{15}{8} \right)^{(k_{max}+1)} > \frac{7}{15} j \quad (\text{A58})$$

So

$$\sum_{\{j_k\}|j} 1 \leq j^{k_{max}} \left(\frac{7}{15} j \right)^{-\frac{k_{max}}{2}} = \left(\frac{15}{7} j \right)^{\frac{k_{max}}{2}} \quad (\text{A59})$$

Using the inequality $k_{max} \leq \frac{\ln((7/15)(j+2))}{\ln(15/8)}$ derived before, we obtain:

$$\sum_{\{j_k\}|j} 1 \leq \exp \left(\ln \left(\frac{15}{7} j \right) \frac{k_{max}}{2} \right) \leq \exp \left(\ln \left(\frac{15}{7} j \right) \frac{\ln((7/15)(j+2))}{2 \ln(15/8)} \right) \quad (\text{A60})$$

Expression on the right-hand side goes as $\exp(\ln^2 j)$, so it is clearly bounded by some power of e^j . To find the specific factor, we bound

$$\ln \sum_{\{j_k\}|j} 1 \leq \ln \left(\frac{15}{7} j \right) \frac{\ln((7/15)(j+2))}{2 \ln(15/8)} = \frac{\ln j \ln(j+2) + \ln \frac{15}{7} \ln \left(1 + \frac{2}{j} \right) - \ln^2 \frac{15}{7}}{2 \ln(15/8)} \leq \quad (\text{A61})$$

$$\leq \frac{\ln^2(j+2) + \ln \frac{15}{7} \ln 3 - \ln^2 \frac{15}{7}}{2 \ln(15/8)} \quad (\text{A62})$$

now we use the property of a function $\ln^2 x \leq 0.6x$ for $x \geq 3$. As an exercise, let's prove it. First note that second derivative of $\ln^2 x$ changes sign at $\ln x = 1$, $x = e$ and is negative for $x > e$, in particular for $x \geq 3$. Next we calculate the values of $\ln^2 x$ and $0.6x$ as well as their derivatives at $x = 3$, which allows us to prove that $\ln^2 x \leq 0.6x$ at least until $x = 7$. Then we evaluate the slopes at $x = 7$ and see that the slope of $\ln^2 x$ is $0.55 < 0.6$. Since the second derivative is always negative, the $\ln^2 x \leq 0.6x$ will be satisfied for all $x > 7$. A numerical check shows that the inequality can be tightened to $\ln^2 x \leq 0.55x$, but we will use the rigorous 0.6:

$$\ln \sum_{\{j_k\}|j} 1 \leq \frac{0.6(j+2) + \ln \frac{15}{7} \ln 3 - \ln^2 \frac{15}{7}}{2 \ln(15/8)} = \frac{0.3j + 0.73}{\ln(15/8)} \quad (\text{A63})$$

Thus

$$\sum_{\{j_k\}|j} 1 \leq (15/8)^{0.3j+0.73} \quad (\text{A64})$$

Now we have all the ingredients for the bound on the X_j in the telescopic sum:

$$\|X_j\| \leq \sum_{\{j_k\}|j} \|X_{\{j_k\}}\| \leq \max_{\{j_k\}|j} \|X_{\{j_k\}}\| \sum_{\{j_k\}|j} 1 \leq \quad (\text{A65})$$

$$\leq (2^{2.2})^j 2^{1.8+0.5S(X_0)} \left(\frac{30}{7}c'\right)^j \chi^{\left(\frac{8}{15}\right)^2 \frac{7}{8}j} (15/8)^{0.3j+0.73} \leq 6 \cdot 2^{0.5S(X)} (24c'\chi^{\frac{56}{225}})^j \|X\| \quad (\text{A66})$$

c. Final step Note that with the bound telescopic sum we can derive a bound for collared $\tilde{X} = \prod_k e^{A_c^{(k)}} X \prod_{k'} e^{-A_c^{(k')}}$, where c is the extra support on each side of X that is allowed to appear in \tilde{X} . To that accord, $A_c^{(k)}$ for every k is chosen to include only terms in A that lie fully within a collar of c around X . Clearly $A_c^{(k)}$ obeys the same bounds as $A^{(k)}$, so in particular the decomposition $\tilde{X} = \sum_j \tilde{X}_j$ obeys the bound (A66). We see that $\prod_k e^{A^{(k)}} X \prod_{k'} e^{-A^{(k')}}$ and \tilde{X} only differ in terms starting from X_{c+1} (because they involve $A_i^{(k)}$'s in regions where \tilde{X} is different). So the difference is bounded by the sum of those:

$$\rightarrow \left\| \prod_k e^{A^{(k)}} X \prod_{k'} e^{-A^{(k')}} - \prod_k e^{A_c^{(k)}} X \prod_{k'} e^{-A_c^{(k')}} \right\| \leq \sum_{j=c+1}^{\infty} \|X_j\| + \|\tilde{X}_j\| \leq 2 \sum_{j=c+1}^{\infty} \|X_j\| \quad (\text{A67})$$

The sum converges:

$$\sum_{j=c+1}^{\infty} \|X_j\| \leq \sum_{j=c+1}^{\infty} 6 \cdot 2^{0.5S(X)} (24c'\chi^{\frac{56}{225}})^j \|X\| = 6 \cdot 2^{0.5S(X)} \frac{(24c'\chi^{\frac{56}{225}})^{c+1}}{1 - 24c'\chi^{\frac{56}{225}}} \|X\| \quad (\text{A68})$$

for $24c'\chi^{\frac{56}{225}} < 1$ which requires very small $\chi \lesssim \frac{10^{-7}}{c^4}$ due to the untight bounds used in the derivation. Note that the original Imbrie's construction also required at least $\chi^{1/20} < 0.5$ which results in $\chi \lesssim 10^{-6}$ — it's not new for such proofs to accumulate large numbers. We can bound the denominator by 2 for $24c'\chi^{\frac{56}{225}} < 0.5$: The sum converges:

$$\sum_{j=c+1}^{\infty} \|X_j\| \leq 12 \cdot 2^{0.5S(X)} (24c'\chi^{\frac{56}{225}})^{c+1} \|X\| \quad (\text{A69})$$

which results in the bound for the collared approximation:

$$\left\| \prod_k e^{A^{(k)}} X \prod_{k'} e^{-A^{(k')}} - \prod_k e^{A_c^{(k)}} X \prod_{k'} e^{-A_c^{(k')}} \right\| \leq 24 \cdot 2^{0.5S(X)} (24c'\chi^{\frac{56}{225}})^{c+1} \|X\| \quad (\text{A70})$$

□

2. Circuit approximation

We would like to have a FDL circuit approximation of a unitary e^A . Consider a simple case of $S(A) = 2$ first: $A = \sum_i A_{i,i+1}$. The Trotter approximation in N steps is defined as:

$$U_T = (e^{\frac{1}{N}A_{\text{even}}} e^{\frac{1}{N}A_{\text{odd}}})^N \quad (\text{A71})$$

where $A_{\text{even}} = \sum_i A_{i,i+1}$. The error is given by the next term in BCH formula while it remains small:

$$e^A = U_T + O(\|[A_{\text{even}}, A_{\text{odd}}]\|/N) = U_T + O(L\|A_{i,i+1}\|^2/N) \quad (\text{A72})$$

Note that the error over whole macroscopic system diverges with size. However, over a small patch the error is small. Let's see how the divergence arises. Here we use big $O(x)$ notation in a sense that $\|e^A - U_T\| \leq Mx$ for some range of x around 0. To get the actual value of constant M in the bound, one way is to expand the following combination of unitaries e^ϵ, e^δ :

$$e^\epsilon e^\delta - e^{\epsilon+\delta} = \frac{[\epsilon, \delta]}{2} + \sum_{j=3}^{\infty} \frac{\epsilon^j + \delta^j - (\epsilon + \delta)^j}{j!} + \epsilon \sum_{j=2}^{\infty} \frac{\delta^j}{j!} + \delta \sum_{j=2}^{\infty} \frac{\epsilon^j}{j!} + \frac{\epsilon}{2} \sum_{j=1}^{\infty} \frac{\delta^j}{j!} + \frac{\delta}{2} \sum_{j=1}^{\infty} \frac{\epsilon^j}{j!} \quad (\text{A73})$$

Taking the norm of the above expression, one arrives to multiple series like $\sum_{j=2}^{\infty} \frac{\|\epsilon\|^j}{j!} = e^{\|\epsilon\|} - 1 - \epsilon$. To upper bound this function one uses a version of Taylor expansion with a mean-value form of the remained:

$$e^x = 1 + x + \frac{e^{\tilde{x}} x^2}{2} \quad \text{where} \quad \tilde{x} \leq x \quad (\text{A74})$$

With that we will be able to derive

$$e^\epsilon e^\delta - e^{\epsilon+\delta} = \gamma, \quad \|\gamma\| \leq M\|[\epsilon, \delta]\| + O(\|\epsilon\|^3, \|\delta\|^3, \|\epsilon\|^2\|\delta\|, \|\epsilon\|\|\delta\|^2) \quad (\text{A75})$$

And the number in front of every term can be obtained explicitly. In fact, the higher order terms are also combinations of commutators ($e^\epsilon e^\delta$ is an element of a Lie group that can be defined free from matrix multiplication), so higher powers of L only appear as $L^p \|A_i\|^{2p}$ or lower, but we will not need it for our purposes here. We now plug in $e^{\frac{1}{N} A_{\text{even}}} e^{\frac{1}{N} A_{\text{odd}}} - e^{\frac{1}{N} A} = \gamma$ where γ is bounded as above into the Trotter formula:

$$\|e^A - U_T\| = (1 + \|\gamma\|)^N - 1 \leq e^{\|\gamma\|N} - 1 \quad (\text{A76})$$

where we have used that all the coefficients in the Taylor expansion of $e^{\|\gamma\|N}$ are strictly bigger than ones in $(1 + \|\gamma\|)^N$. Now we can again use the mean-value form of the remainder of the Taylor expansion to bound the above with explicit constant as long as $e^{\|\gamma\|N}$ is not too big:

$$\|e^A - U_T\| \leq O(\|[A_{\text{even}}, A_{\text{odd}}]\|/N) = O(L\|A_{i,i+1}\|^2/N) \quad (\text{A77})$$

We will stick with the big O notation for convenience. Let's prove that for an operator X supported over $S(X)$ sites the conjugation with e^A is simulated by our Trotter approximation with system-size independent error:

Lemma A.3. for $A = \sum_i A_{i,i+1}$

$$e^A X e^{-A} = U_T X U_T^\dagger + O\left(\|X\| \frac{(S(X) + 4N)}{N} \|A_{i,i+1}\|^2\right) \quad (\text{A78})$$

We see that the system size dependence drops out, also $U_T X U_T^\dagger$ is an operator supported on $S(X) + 4N$ sites, and expressed by a tensor network of size $(S(X) + 4N) \times 4N$. The terms outside the causal cone of X just cancel.

Proof. First we use the bounds like those in Lemma A.1 for $e^A X e^{-A}$:

$$e^A X e^{-A} = \sum_j X_j, \quad S(X_j) = S(X) + 2j, \quad \|X_j\| \leq 2^{0.5S(X)} (4\|A_{i,i+1}\|)^j \|X\| \quad (\text{A79})$$

Then we use the approximation by collared evolution like the one in Theorem A.2 (the final part of the proof)

$$\|e^A X e^{-A} - e^{A_c} X e^{-A_c}\| \leq 2 \sum_{j=c+1}^{\infty} \|X_j\| \leq 4 \cdot 2^{0.5S(X)} (4\|A_{i,i+1}\|)^{c+1} \|X\| \quad (\text{A80})$$

Here c is the number of sites we are to include in the "collar" of X . A_c are the terms in A supported inside the collared region. The problem reduced to finding an FDL approximation to e^{A_c} supported on $S(X) + 2c$ - a finite number of spins. For that purpose, Trotter approximation is used. The total precision defined as:

$$\|e^A X e^{-A} - e_T^{A_c} X e_T^{-A_c}\| \leq 4 \cdot 2^{0.5S(X)} (4\|A_{i,i+1}\|)^{c+1} \|X\| + O((2c + S(X))\|A_{i,i+1}\|^2 \|X\|/N) \quad (\text{A81})$$

Note that if we were Trotter approximating e^A instead of the collared one, we still get the same result $e_T^A X e_T^{-A} = e_T^{A_c} X e_T^{-A_c}$ by cancellations of the spare terms as long as $c \geq 2N$. So there exists an FDL circuit e_T^A over the whole system that approximates e^A locally in every patch.

Plugging in the collar as small as $c = 2$, we find the first term $\sim (4\|A_{i,i+1}\|)^3$ to be negligible (we assume $2^{0.5S(X)}$ is suppressed by powers of $\|A_{i,i+1}\|$ for big O bound to make sense). The only allowed N for $c = 2$ for the total circuit to reduce to collared circuit is $N = 1$. The error is now:

$$\|e^A X e^{-A} - e_T^{A_c} X e_T^{-A_c}\| \leq O((4 + S(X))\|A_{i,i+1}\|^2 \|X\|) \quad (\text{A82})$$

More generally, we can take bigger $c = 2N$, and the first term will always be suppressed as $\sim (4\|A_{i,i+1}\|)^{2N+1}$, which allows to neglect it compared to the second term:

$$\|e^A X e^{-A} - e_T^{A_c} X e_T^{-A_c}\| \leq O((4N + S(X))\|A_{i,i+1}\|^2 \|X\|/N) \quad (\text{A83})$$

□

Note that the precision never really gets better than $O(\|A_{i,i+1}\|^2 \|X\|)$. That's concerning because our original collar $c = 2$ approximation already had $\sim (4\|A_{i,i+1}\|)^3$ precision. So if we want to approximate that well, we need to resort to a higher order Trotter scheme (instead of $e^{\frac{1}{N}A_{\text{even}}} e^{\frac{1}{N}A_{\text{odd}}}$ it uses a more complicated step). Alternatively, since the dimension of the space we act upon ($d = 2^{S(X)+4}$ for $c = 2$) is relatively small, we can instead use the Solovay-Kitaev theorem¹⁷. For this particular case it can be phrased as follows: to approximate a unitary on a d -dimensional Hilbert space to a precision ϵ , one needs $n = O\left(d^a \ln^b \frac{1}{\epsilon}\right)$ 2-qubit gates acting on nearest neighbors in 1d. Here a, b are some constant numbers. So to achieve precision $\epsilon = \|A_{i,i+1}\|^{c+1}$, one would need $n = O(c2^{ac})$ gates, or $O(2^{ac})$ depth. The number of gates in the c 'th order of Trotter step also scales exponentially with c , giving the same $\|A_{i,i+1}\|^{c+1}$ precision. So whichever method wins depends on the constant a and the scaling of the size of Trotter step.

Now let's apply this logic to the Imbrie circuit $\prod_{k=1}^{\infty} e^{A^{(k)}}$. We will see that as early as for approximation up to $k = 2$ the first order Trotter scheme becomes useless. Indeed, the $S(A_k) = \text{floor}(\frac{8}{7}L_{k+1})$, so $S(A_1) = 4$, $S(A_2) = 7$, $S(A_3) = 14$. The telescopic sum structure $X = X_j$ allows us to drop higher order steps, since step j only uses the terms with $S(A_k) - 1 \leq j$, and make an error given by Theorem A.2:

$$\left\| \prod_k e^{A^{(k)}} X \prod_{k'} e^{-A^{(k')}} - \prod_k e^{A_c^{(k)}} X \prod_{k'} e^{-A_c^{(k')}} \right\| \leq 24 \cdot 2^{0.5S(X)} (24c' \chi^{\frac{56}{225}})^{c+1} \|X\| \quad (\text{A84})$$

For $c < 6$, $S(A_2) - 1 = 6 > c$, so we only include the 1st order term in the product and obtain the correct $X_1 \dots X_6$:

$$\left\| \prod_k e^{A^{(k)}} X \prod_{k'} e^{-A^{(k')}} - e^{A_c^{(1)}} X e^{-A_c^{(1)}} \right\| \leq 24 \cdot 2^{0.5S(X)} (24c' \chi^{\frac{56}{225}})^{c+1} \|X\| \quad (\text{A85})$$

Now, the $A_i^{(1)}$ has support 4, but actually the first step of perturbation theory has support 3, just our definition overestimates it for the sake of consistent form over all k . So it's enough to group spins into pairs before applying the Trotter: $A_{\mu,\mu+1} = A_i^{(1)} + A_{i+1}^{(1)}$ and $i = 2\mu - 1$. As a result we get an approximation $e_T^{-A_c^{(1)}}$ by 4-local gates bounded using Lemma A.3 as:

$$\|e^{A_c^{(1)}} X e^{-A_c^{(1)}} - e_T^{A_c^{(1)}} X e_T^{-A_c^{(1)}}\| \leq O((4N + S(X))\|A_{\mu,\mu+1}\|^2 \|X\|/N) \leq O((4N + S(X))(c'\chi)^2 \|X\|/N) \quad (\text{A86})$$

and due to the grouping the condition of one circuit for the whole system is $4N \leq c$. It can only be satisfied for $c \geq 4$, and $N = \text{floor}(c/4) = 1$ always:

$$\left\| \prod_k e^{A^{(k)}} X \prod_{k'} e^{-A^{(k')}} - e_T^{A_c^{(1)}} X e_T^{-A_c^{(1)}} \right\| \leq 24 \cdot 2^{0.5S(X)} (24c' \chi^{\frac{56}{225}})^{c+1} \|X\| + O((4 + S(X))(c'\chi)^2 \|X\|) \quad (\text{A87})$$

Comparing the powers, we note that $\frac{56(c+1)}{225} > 2$ for $c \geq 4$, so the first term can be omitted for sufficiently small χ . We found that Trotter-approximating the first step of perturbation theory leads to an error

$$\left\| \prod_k e^{A^{(k)}} X \prod_{k'} e^{-A^{(k')}} - e_T^{A_c^{(1)}} X e_T^{-A_c^{(1)}} \right\| \leq O((4 + S(X))(c'\chi)^2 \|X\|) \quad (\text{A88})$$

everywhere except for rare resonances. One can chose depth 2 and 4-local gates, or depth 3 and 3-local gates (the second application of Trotter bound leads to the same scaling of error). To go down to 2-local gates, one may apply Solovay-Kitaev to the 3-local gates, blowing up the depth by $O\left(8^a \ln^b \frac{1}{\epsilon}\right) = O(|\ln \chi|^b)$ factor. Note that even though for a uniform deep circuit the causal structure would make $e_T^{A_c^{(1)}} X e_T^{-A_c^{(1)}}$ depend on $O(|\ln \chi|^b)$ collar, the long patches of gates arising from Solovay-Kitaev are arranged in a checkerboard pattern, so the true causal cone remains within c .

To conclude, let us study the relationship between a state $\prod_k e^{A^{(k)}} X |prod\rangle$ and $e_T^{A_c^{(1)}} |prod\rangle$. Any observable X on $S(X)$ sites is approximated by measuring it with $e_T^A |0\rangle$ with error as in Eq. (A88). The latter measurement is represented by a system-size independent tensor network due to the cancellations. In particular, we can measure a matrix element of the density matrix over those $S(X)$ sites. The difference in density matrices over $S(X)$ sites is bounded element-wise:

$$|(\rho - \rho_{appr})_{\alpha\beta}| \leq O((4 + S(X))(c'\chi)^2 \|X\|) \quad (\text{A89})$$

We can also show the operator norm bound:

$$\|\rho - \rho_{appr}\| \leq O((4 + S(X))(c'\chi)^2 \|X\|) \quad (\text{A90})$$

here $\rho = \text{tr}_{\overline{S(X)}} \prod_k e^{A^{(k)}} |prod\rangle \langle prod| \prod_{k'} e^{-A^{(k')}}$ and $\rho_{appr} = \text{tr}_{\overline{S(X)}} e_T^{A_c^{(1)}} |prod\rangle \langle prod| e_T^{-A_c^{(1)}}$. $|\prod\rangle$ is a product state in z -basis. To prove the norm bound, consider the eigenbasis of $\rho - \rho_{appr}$. We can make X to be measuring matrix elements in this basis. In particular, the maximum eigenvalue of $\rho - \rho_{appr}$ will not be bigger than the difference $|\langle X \rangle - \langle X \rangle_{appr}| \leq O((4 + S(X))(c'\chi)^2 \|X\|)$. Note that for the bound to be correct, the size for causal cone of X should be $S(X) + 2c$ and $c \geq 4$, and it does not improve the bound to have $c > 5$.

To go to a better precision than $(c'\chi)^2$, specifically χ^p , the depth required is exponential in p until we allow some threshold size $S(X) + 2c \sim p$ of gates in our circuit. Indeed, even for $A = \sum_i A_{i,i+1}$ above there isn't a subexponential way to achieve high precision using methods described so far. The threshold is reached when that $\sum_{j=1}^c X_j$ needed to approximate to precision χ^p is contained entirely within one gate, so we can do $\prod_k^{k_{max}} e^{A_c^{(k)}}$ in one step. This does not allow us to have a total circuit though - so methods described so far always require exponential in p size of the circuit if one wants to have one circuit working for operator in any place of the system that reduces to a collared one via cancellations. There might be a method of getting a total circuit of $\sim p$ local gates and subexponential depth via splitting $e^A = (e^{A_L} \otimes e^{A_R}) e^{G_{LR}}$ where L, R are half-chains one the left and right of a specific cut. The G_{LR} defined in this way does not commute with A_L, A_R , so the error above may be $\sim \|A_R\|$ - divergent with the system size. In fact it's not! The G_{LR} has a collar bound/telescopic sum around the cut that can be found in¹⁶, but we do not pursue the construction of precise subexponential depth circuit here.

3. Area law

We will prove the following:

Corollary A.3.1. *The state $\prod_{k=1} e^{A^{(k)}} |prod\rangle$ obeys the area law: for B - the half of the system, the entanglement entropy $S(\rho_B) \leq 650c'\chi$. Here $|prod\rangle$ is a product state in the z -basis.*

We will also discuss the tail of the distribution of entanglement entropies in the presence of resonances.

We turn to the paper¹⁵ which derives a bound on entanglement production $S(t) - S(0)$ under the evolution e^{iHt} . In our case, we apply the bound for every step using $iHt = A^k$. We set $t = 1$.

We use the bound that is given on the last page of¹⁵. The interaction between two parts in $H = -iA^k$ can be written as a Shmidt decomposition $H_{LR}^{(k)} = \sum_n r_n J_L^n \otimes J_R^n$, where we have freedom to set $\|J_L\| = 1, \|J_R\| = 1$. The A_i^k are bounded by $\|A_i^{(k)}\| \leq c'\chi^{\text{ceil} L_{k-1}}$ and there are $S(A^{(k)}) - 1 = \text{floor}(\frac{8}{7}L_{k+1} - 1)$ of them across the cut, thus the interaction is bounded as $\|H_{LR}^{(k)}\| \leq c'(S(A^{(k)}) - 1)\chi^{\text{ceil} L_{k-1}}$. We can decompose each A_i^k across the cut into $4^{S(A_{i,L}^k)}$ terms — this is how much basis operators are there on one side of the cut. Each term in the decomposition is bounded as the whole operator that's decomposed $\|A_i^{(k)}\|$:

$$\|r_n\| \leq c'\chi^{\text{ceil} L_{k-1}} \quad (\text{A91})$$

There are a total of $\sum_{S(A_{i,L}^k)=1}^{S(A^{(k)})/2} 4^{S(A_{i,L}^k)} \leq \frac{4}{3} 2^{S(A^{(k)})}$ terms. The entanglement produced over time $t = 1$ will be

$$S(t) - S(0) \leq c^* \sum_i \|r_i\| \leq c^* \frac{4}{3} 2^{S(A^{(k)})} c'\chi^{\text{ceil} L_{k-1}} \quad (\text{A92})$$

The coefficient $c^* \approx 1.9$ is universal and is found in¹⁴. The total entanglement is the sum of entanglement produced by each order k of the circuit. The total bound on entanglement is converging:

$$\delta S_{tot} \leq \sum_k c^* \frac{4}{3} c' 2^{\text{floor}(\frac{8}{7}L_{k+1})} \chi^{\text{ceil}L_{k-1}} \leq \quad (\text{A93})$$

$$\leq c^* \frac{4}{3} c' 2^{\text{floor}(\frac{8}{7}L_2)} \chi^{\text{ceil}L_0} \sum_{j=0}^{\infty} \left(2^{\text{floor}(\frac{8}{7}L_3) - \text{floor}(\frac{8}{7}L_2)} \chi^{\text{ceil}L_1 - \text{ceil}L_0} \right)^j \leq \quad (\text{A94})$$

$$\leq \frac{c^* \frac{4}{3} c' 2^7 \chi}{1 - 2^7 \chi} \leq c^* \frac{4}{3} c' 2^8 \chi \leq 650 c' \chi \quad (\text{A95})$$

where we replaced the sum of $\exp(L_k) = \exp((15/8)^k)$ by a geometric series with a factor such that every term upper-bounds the corresponding term in the original sum, and $\chi < 1/2^8$ has been used to get rid of the denominator. Thus the Imbrie states without resonances satisfy area law.

Every resonance is an extra local rotation over region of size n_r that does not have to be small. If it happens that our cut passes through the resonance, we need to add the entanglement created by this rotation to the total entanglement bound. A rotation over n_r sites can create $\Delta S(n_r) \leq n_r \ln 2$ entanglement. Naive expectation is $\leq (n_r/2) \ln 2$ which is true for a unitary acting on a whole system of size n_r . But in fact having a bigger system we can create two times more entanglement with the rotation on n_r sites! Indeed, the bound is saturated by a swap gate, that turns a state with $n_r/2$ Bell pairs on every side of the cut into a state where every Bell pair crosses the cut. We believe that this is an overestimation and only two levels are mixed at a typical resonance, producing $\sim \ln 2$ entanglement.

After k_1 when the first resonance appears across the cut, there may be other, bigger resonances containing the first one. Since a cut is placed randomly in the system, denote the probability that the biggest resonance across the cut has size n_r by $P(n_r)$. It is exactly the same quantity as the probability of a site to be a part of a connected cluster of resonances of size n_r . Imbrie shows that this probability decays faster than any polynomial. Thus the expectation value

$$\sum_{n_r} n_r P(n_r) \leq \text{const} \quad (\text{A96})$$

Moreover, from investigating the first step of the construction we know that the first term has $n_r = 3$ and $P(n_r = 3) = \epsilon$ (with some coefficient depending on the probability distribution of disorder, in our case 1) and then it keeps decaying faster than any polynomial, so the leading contribution to the expectation value is the first term:

$$\sum_{n_r} n_r P(n_r) \leq 4\epsilon \quad (\text{A97})$$

Here the first factor of 2 is the bound on entanglement instead of 3, as one of the sides has one site, and the second factor of 2 absorbs the later terms in the sum, which can be done for sufficiently small ϵ . Thus the total bound on the expectation value of the entanglement entropy is:

$$\overline{S_{tot}} \leq 650 c' \chi + 4\epsilon \ln 2 \quad (\text{A98})$$

We note that the first term $\gamma^{0.95}$ is smaller than the second term $\gamma^{0.05}$ — so the main contribution to average entanglement is from rare distributions of disorder when resonances cross the cut. We also note that in full construction by Imbrie, the perturbative parts of the circuit $e^{A^{(k)}}$ get adjusted in the presence of resonances as the resonant spins are grouped into metaspins. The entanglement produced by resulting unitaries is still bounded, so we dropped this complication from our simplified circuit.

Consider rare cuts that may have entanglement bigger than $2\ln 2$, by passing through the big resonant regions. Assume for simplicity exponential distribution of entanglement as a discrete random variable $P(S = 2x \ln 2) = \epsilon^x$ for x the half of the size of the resonance. The biggest entanglement across the cut in a given system determined by:

$$LP(S_{max}) = L\epsilon^{-S_{max}/2\ln 2} \approx 1 \quad (\text{A99})$$

$$S_{max} \approx 2\ln 2 \frac{\ln L}{(-\ln \epsilon)} \quad (\text{A100})$$

To derive maximal entanglement more carefully one needs to use the true statistics of resonances from Imbrie's paper, summarized in Postulate 4, that results in a superpolynomially decaying $P(S)$ anyway. It only adjusts $\ln L \rightarrow e^{\sqrt{\ln L}}$ - a difference between the two will be almost impossible to detect.

4. LRB

Consider

$$A(t) = e^{iHt} A e^{-iHt} = U^\dagger e^{iH_z t} U A U^\dagger e^{-iH_z t} U = \sum_r U^\dagger e^{iH_z t} A_r e^{-iH_z t} U \quad (\text{A101})$$

where H is the Hamiltonian of our spin chain and $U = \prod_{k=1} e^{A^{(k)}}$ is the unitary that diagonalizes it. According to Theorem A.2, A_r supported on $S(A) + 2r$ sites and is bounded by $6 \cdot 2^{0.5S(A)} \alpha^r \|A\|$ with $\alpha = 24c' \chi^{\frac{56}{225}}$. Now the terms in $e^{iH_z t}$ that do not overlap with A_r commute through, so

$$e^{iH_z t} A_r e^{-iH_z t} = e^{iH_z^r t} A_r e^{-iH_z^r t} \quad (\text{A102})$$

where $H_z^r = \sum_m H_z^{r+m}$ and the terms are bounded as $36\alpha^m$ (using $H = \sum_i H_i$, $\|H_i\| \leq 2 + \gamma \leq 3$ and $S(H_i) = 2$). Now we find m^* such that $\alpha^{m^*} t = 1$. Define

$$Q_{r+m^*+p} = e^{i \sum_m^{m^*+p} H_z^{r+m} t} A_r e^{-i \sum_m^{m^*+p} H_z^{r+m} t} - e^{i \sum_m^{m^*+p-1} H_z^{r+m} t} A_r e^{-i \sum_m^{m^*+p-1} H_z^{r+m} t} \quad (\text{A103})$$

and $Q_{r+m^*} = e^{i \sum_m^{m^*} H_z^{r+m} t} A_r e^{-i \sum_m^{m^*} H_z^{r+m} t}$. The higher-order Q 's are bounded

$$\|Q_{r+m^*+p}\| \leq \|e^{iH_z^{r+m^*+p} t} - 1\| \|A_r\| \leq \|A\| c^* \alpha^p \quad (\text{A104})$$

Where c^* is a new constant coefficient. We found that the telescopic sum for $e^{iH_z^r t} A_r e^{-iH_z^r t}$ has the form

$$e^{iH_z^r t} A_r e^{-iH_z^r t} = \sum_p Q_{r+m^*+p}, \quad \|Q_{r+m^*+p}\| \leq \|A_r\| c^* \alpha^p \quad (\text{A105})$$

After plugging it in the decomposition of $A(t)$:

$$A(t) = \sum_r \sum_p U^\dagger Q_{r+m^*+p} e^{-iH_z t} U \quad (\text{A106})$$

The last conjugation with the Imbrie circuit produces one more telescopic sum. Resumming all three telescopic sums, we get:

$$A(t) = \sum_q Q'_{m^*}, \quad \|Q'_{m^*+q}\| \leq \|A_r\| \frac{1}{2} c^{**} \alpha^q \quad (\text{A107})$$

A new constant and a factor $\frac{1}{2}$ was introduced for convenience. Now a local operator B distance d_{AB} from A will only overlap with $m^* + q \geq d_{AB}$ terms in the telescopic sum. Resumming once more, we obtain:

$$\|[A(t), B]\| \leq \|A\| \|B\| c^{**} \alpha^{d_{A,B}-m^*} = \|A\| \|B\| c^{**} \alpha^{d_{A,B} t} \quad (\text{A108})$$

That concludes the proof.

Appendix B: Interpretation of Imbrie's proof

1. Bound on generators

Here we discuss the bound on the local generator $A_i^{(k)}$ of the perturbative step $e^{\sum_i A_i^{(k)}}$ of Imbrie circuit. Here i is the spatial index labeling sites of the chain, and k is the order of the perturbation theory. Each $A_i^{(k)}$ is acting on $\text{floor}(\frac{15}{7} L_k)$ spins, where the radius of locality $L_k = (\frac{7}{8})^k$ is growing with k . We have presented a bound

$$\|A_i^{(k)}\| \leq c' \chi^{\text{ceil} L_{k-1}} \quad (\text{B1})$$

Imbrie works with a different decomposition: instead of $A^{(k)} = \sum_i A_i^{(k)}$ he uses $A^{(k)} = \sum_g A^{(k)}(g)$, where the sum iterator g will be explained below. On p. 27 of the arXiv version of the paper¹, Imbrie presents a combinatorial argument that $\sum_{g|i} A^{(k)}(g)$ of terms contributing to $A_i^{(k)}$ converges, and moreover the Bound (B1) holds, however the Bound is not explicitly presented.

Let's define the path formalism used by Imbrie in the following toy examples to facilitate the understanding of the argument on p. 27.

a. *Toy example 1.* Consider an operator J defined as:

$$J = \sum_g J(g) \quad (\text{B2})$$

The sum is over paths $g = \{i_1 \dots i_{|g|}\}$ in the space of spin flips. Each i_n is a site on a lattice, and the length of the string of those sites g is denoted $|g|$. In the \sum_g , it is implied that $|g| > 1$. There's one condition on the string g : each new spin flip has to be within 1 lattice spacing from one of the old ones:

$$\forall n > 1 \exists n' < n : |i_n - i_{n'}| \leq 1 \quad (\text{B3})$$

The operators $J(g)$ are supported on the smallest consecutive region containing all spin flips in g plus a 1 spin collar on the left and on the right. Consider the following way to bound the terms in J :

$$|J_{\sigma\sigma'}(g)| \leq \frac{\delta^{|g|}}{C(|g|)} \quad (\text{B4})$$

where $C(x)$ is a combinatorial factor to be determined and $\delta \ll 1$. We'd like to see what $C(x)$ do we need to be able to translate this bound to a more traditional norm bound. Consider a decomposition of J into "local" terms centered around sites of the lattice:

$$J = \sum_i J_i, \quad J_i = \sum_{g|i_1=i} J(g) \quad (\text{B5})$$

where J_i contains only graphs starting at i_1 . Let's show that a norm bound on J_i follows from a bound on $J_{\sigma\sigma'}(g)$:

$$\|J_i\| \leq \sum_{g|i_1=i} \|J(g)\| \quad (\text{B6})$$

Note that any operator $J(g)$ supported on $\cup g \cup c$ where c is a 1-spin collar can be written in an operator basis with one basis vector per matrix element :

$$J(g) = \sum_{\sigma\sigma'} J_{\sigma\sigma'}(g) |\sigma\rangle\langle\sigma'| \quad (\text{B7})$$

and since $J(g)$ acts as an identity outside $\cup g \cup c$, and doesn't flip c , the number of terms in the sum is $4^{|\cup g|} 2^{|c|} \leq 2^{2|g|+c}$:

$$\sum_{g|i_1=i} \|J(g)\| \leq \sum_{g|i_1=i} \sum_{\sigma\sigma'} J_{\sigma\sigma'}(g) \leq \sum_{g|i_1=i} 2^{2|g|+C} \frac{\delta^{|g|}}{C(|g|)} \quad (\text{B8})$$

We can now split the sum into terms corresponding to each $|g|$:

$$\|J_i\| \leq \sum_{x=1}^{\infty} 2^{2x+C} \frac{\delta^x}{C(x)} \sum_{g|i_1=i, |g|=x} 1 \quad (\text{B9})$$

A simple upper bound on the number of elements in the sum $\sum_{g|i_1=i, |g|=x} 1$ is if a flip i_n can be chosen from $2n-1$ sites centered at i_1 . For $|g|=x$ flips, it gives $(2x-1)!! = (2x-1)(2x-3)\dots$.

$$\|J_i\| \leq \sum_{x=1}^{\infty} 2^{2x+C} \delta^x \frac{(2x-1)!!}{C(x)} \quad (\text{B10})$$

So a choice of $C(x) = (2x-1)!!$ will cancel this factor. What remains is a geometric series which is bounded for $4\delta < 1/2$:

$$\|J_i\| \leq 2^{C+1} (4\delta) \quad (\text{B11})$$

we have arrived at a desired bound.

b. Toy example 2. We will repeat the above construction with a minor modification: in the

$$J = \sum_g J(g) \quad (\text{B12})$$

the \sum_g is now taken over g such that $|g| > L_1$. The bound on $J_{\sigma\sigma'}(g)$ and $C(x) = (2x-1)!!$ are the chosen the same. All the steps go through in exactly the same way, until we arrive to:

$$\|J_i\| \leq \sum_{x=L_1}^{\infty} 2^{2x+C} \delta^x \frac{(2x-1)!!}{C(x)} = \sum_{x=L_1}^{\infty} 2^{2x+C} \delta^x \leq 2^{C+1} (4\delta)^{L_1} \quad (\text{B13})$$

we see that any level is bounded by the smallest allowed flip.

c. Toy example 3. We will now introduce an extra layer of structure to match Imbrie's construction. Consider

$$J = \sum_G J(G) \quad (\text{B14})$$

where $G = \{g_1, g_2 \dots g_{|G|}\}$ a string of spin flip paths g_n defined in examples before. The constraint on the paths in G is now:

$$\forall n > 1 \exists n' < n : d(g_n, g_{n'}) \leq 1 \quad (\text{B15})$$

where $d(g_1, g_2)$ is the distance in lattice spacings between the two nearest spins in g_1, g_2 .

The support of $J(G)$ is in the smallest consecutive region containing all spin flips of underlying paths of G plus some constant collar c . Surprisingly, it is enough to bound

$$|J_{\sigma\sigma'}(G)| \leq \frac{1}{C(|G|)} \prod_{g \in G} \frac{\delta^{|g|}}{C(|g|)} \quad (\text{B16})$$

where $C(x) = (2x-1)!!$ again. Indeed, for a "local" term

$$J_i = \sum_{G|i=i_1 \in g_1 \in G} J(G) \quad (\text{B17})$$

the norm bound

$$\|J_i\| \leq \sum_{G|i=i_1 \in g_1 \in G} \|J(G)\| \leq \sum_{G|i=i_1 \in g_1 \in G} \frac{4^{\sum |g|} 2^c}{C(|G|)} \prod_{g \in G} \frac{\delta^{|g|}}{C(|g|)} \quad (\text{B18})$$

The sum over G can be split into sums over paths with fixed values of $|G|$ and all $|g_m|$:

$$\|J_i\| \leq \sum_{|G|, |g_m|} \frac{4^{\sum |g_m|} 2^c}{C(|G|)} \prod_m \frac{\delta^{|g_m|}}{C(|g_m|)} \sum_{G|i=i_1 \in g_1 \in G, |G|, |g_m|} 1 \quad (\text{B19})$$

so we need to do the counting again. We first count the small graphs as before, and then count how many ways are there to place $|G|$ small graphs of sizes $|g_m|$ so that each new one is a neighbor to one of the old ones. We get the following bound:

$$\sum_{G|i=i_1 \in g_1 \in G, |G|, |g_m|} 1 \leq \quad (\text{B20})$$

$$\leq \prod_m C(|g_m|) |g_1| (|g_1| + 2|g_2|) (|g_1| + 2|g_2| + 2|g_3|) \dots (2 \sum_m |g_m| - |g_1|) \quad (\text{B21})$$

or if we plug it in:

$$\|J_i\| \leq \quad (\text{B22})$$

$$\leq \sum_{|G|, |g_m|} \frac{4^{\sum |g_m|} 2^c \delta^{\sum_m |g_m|} |g_1| (|g_1| + 2|g_2|) (|g_1| + 2|g_2| + 2|g_3|) \dots (2 \sum_m |g_m| - |g_1|)}{C(|G|)} \quad (\text{B23})$$

Or if we separate sums over $|G|, \sum_m |g_m| = X, |g_m| = x_m$:

$$\|J_i\| \leq \sum_{|G|} \sum_{X \geq |G|} 4^X 2^c \delta^X \sum_{x_m | \sum_m x_m = X} \frac{x_1(x_1 + 2x_2)(x_1 + 2x_2 + 2x_3) \dots (2X - x_1)}{C(|G|)} \quad (\text{B24})$$

We see that the smallest power of δ is given by $\sum_m |g_m| = |G|$, in which case the $C(|G|)$ exactly cancels its numerator:

$$\|J_i\| \leq \sum_{|G|} 4^{|G|} 2^c \delta^{|G|} F(|G|, \delta) \quad (\text{B25})$$

where

$$F(|G|, \delta) = 1 + \quad (\text{B26})$$

$$+ \sum_{X > |G|} 4^{X-|G|} \delta^{X-|G|} \sum_{x_m | \sum_m x_m = X} \frac{x_1(x_1 + 2x_2)(x_1 + 2x_2 + 2x_3) \dots (2X - x_1)}{C(|G|)} \quad (\text{B27})$$

we want to show that $f(|G|, \delta)$ can be bounded by a constant for sufficiently small δ . It is a series in terms of $n = X - |G|$. $F = \sum_{n=0}^{\infty} F_n$ and:

$$F_n = (4\delta)^n \sum_{x_m | \sum_m x_m = |G|+n} \frac{x_1(x_1 + 2x_2)(x_1 + 2x_2 + 2x_3) \dots (2(|G| + n) - x_1)}{C(|G|)} \quad (\text{B28})$$

There are always $|G|$ terms x_m . One can already see that

$$\sum_{x_m | \sum_m x_m = |G|+n} \frac{x_1(x_1 + 2x_2)(x_1 + 2x_2 + 2x_3) \dots (2(|G| + n) - x_1)}{C(|G|)} \leq (|G| + n)^{2|G|} \quad (\text{B29})$$

so we only get a polynomial prefactor in front of δ^n and the series converges. What we want to show is that it converges to something that goes to ∞ only exponentially in $|G|$, so we can suppress it by $(4\delta)^{|G|}$. Clearly the naive bound $(|G| + n)^{2|G|}$ where dropped $C(|G|)$ altogether gives a factorial-like scaling even for the first term in the sum that we know is 1. We know that number of terms in the sum for each n is in fact given by $C_{|G|+n}^{|G|} < 2^{|G|+n}$, so what's left is to bound a single term in the sum:

$$\frac{x_1(x_1 + 2x_2)(x_1 + 2x_2 + 2x_3) \dots (2(|G| + n) - x_1)}{C(|G|)} \leq \quad (\text{B30})$$

$$\leq \left(\frac{(2(|G| + n) - 1)}{2|G| - 1} \right)^{|G|} = \left(1 + \frac{2n}{2|G| - 1} \right)^{|G|} \quad (\text{B31})$$

these terms summed with $4\delta^n$ can only lead to logarithmic corrections $(1/(-\ln\delta))^{|G|}$ to the $F(|G|, \delta)$. Putting it all together:

$$F_n \leq (4\delta)^n 2^{|G|+n} \left(1 + \frac{2n}{2|G| - 1} \right)^{|G|} \quad (\text{B32})$$

There is no factorial in $|G|$, so:

$$F(|G|, \delta) = \sum_{n=0}^{\infty} F_n \leq c' = O(1) \quad (\text{B33})$$

dropping the logarithmic corrections, and

$$\|J_i\| \leq \sum_{|G|} c' 2^c (4\delta)^{|G|} \leq c' 2^{c+1} (4\delta) \quad (\text{B34})$$

Each term in the sum is bounded by our choice of $C(|G|)$. We can also consider long paths of $|G| > L_2$ like in the toy example 2 and get $c' 2^{c+1} (4\delta)^{L_2}$.

d. Actual construction The structure of J_i that we described is present in the generators A_i of steps 1 and 2 of Imbrie's construction, with several differences. One difference is that in the original bound on $A_{\sigma\sigma'}(g)$ the γ/ϵ stands in place of δ that we used here. Another difference with J_i is that the paths for A are bounded in length: from below which leads to a higher power of γ/ϵ and from above by a growing scale L_{k+1} , which gives truly local generators, but does not in any way affect the derivation above. Also, the paths in $g^{(k)}$ do not have to touch on a chain - instead, a collar is allowed, which is then lifted to fill in the gap with new graphs as described on pp. 45-46 of paper¹. We discuss how the collar appears in the Subsection B 2.

Finally, in Imbrie's paper $C(x) \neq (2x-1)!! \sim (x!)^{1/2}$, instead it's an arbitrary power $C(x) \sim (x!)^{2/9}$ (the graph factorial is defined on p.23 in the same way as our toy examples). The combinatorial factors are still controlled, but they require a more careful way to bound \sum_g and extra tricks like resumming the long paths as a subroutine, as described on p. 27 of paper¹. We note that in the earlier arXiv version of the paper, the $n!^{1/2}$ was used as a combinatorial factor.

Here we repeat the argument on p.27. The construction results in $H_{pert} = \sum_g J(g)$. Then $\langle \sigma | H_{pert} | \sigma' \rangle = \sum_g J_{\sigma\sigma'}(g)$ where the terms in the sum have a bound proven in¹ for them. We'd like to prove that the sum is bounded as well. For that we do a transformation $\sum_g |J_{\sigma\sigma'}(g)| = \sum_g |J_{\sigma\sigma'}(g)| c(g) c^{-1}(g) \leq \max_g J_{\sigma\sigma'}(g) |c(g)| \sum_g c^{-1}(g)$ where $c(g)$ is some arbitrary function of the graph g . If $c(g)$ is chosen in such a way that $\sum_g c(g)^{-1} \leq 1$, then $\sum_g |J_{\sigma\sigma'}(g)| \leq \max_g J_{\sigma\sigma'}(g) |c(g)|$. Imbrie chooses $c(g)$ in such a way that $\max_g J_{\sigma\sigma'}(g) |c(g)| \leq (a\chi)^{\min|g|}$ with some constant a , and proceeds to show that $\sum_g c(g)^{-1} \leq 1$ for that choice. Since we consider $J_{\sigma\sigma'}(g)$, it is nonzero only for g 's that connect σ and σ' , the $\min|g|$ is at least the length of the interval containing the difference $\sigma - \sigma'$. Similar process can be applied to A , where $\min|g|$ is the minimal remaining order $\text{ceil} L_{k-1}$.

2. Bound on support

Imbrie's diagonalization proceeds in steps labeled by $k = 1, 2, 3, \dots$. At each step a unitary rotation $O_r e^{A^{(k)}}$ is applied to the Hamiltonian, where O_r is an operator acting spatially only near resonances, and as identity everywhere else. The generator of non-resonant part of the circuit is given as a sum of local terms $A^{(k)} = \sum_g A(g)$, where g is the data in the format of the generalized path, containing, in particular, the locations of spin flips of $A(g)$. The number of spin flips is upper bounded by the order of the generalized path $|g|$ — the power of the magnetic field involved in this term (some flips can act on the same location and "cancel" each other). At step k , that order is $L_{k-1} \leq |g| < L_k$, where $L_k = (15/8)^k$ (see, for example, p.24 of the new arXiv version). One may group terms into $A_i^{(k)} = \sum_{g: i \in \text{supp} A(g)} A(g)$ where the first site on which $A(g)$ acts nontrivially is i . We demonstrated in Subsection B 1 how $\|A_i^{(k)}\|$ constructed this way can be bounded knowing the bound on $A(g)$. Here we'll discuss the support of $A(g)$ and $A_i^{(k)}$.

The construction has a separate term $A(g)^{(k)}$ for each matrix element $A(g)^{(k)} \sim |\sigma\rangle\langle\sigma'|$, given by $A(g)^{(k)} = \frac{J(g)^{(k-1)}}{E_{\sigma}^{(k-1)} - E_{\sigma'}^{(k-1)}}$. Here $J^{(k-1)}$ are the terms appearing after $k-1$ previous rotations of H . From the point of view of support, $J(g)^{(k-1)}$ combines support of the lower order paths that are part of g data, while $E_{\sigma}^{(k-1)}$ is calculated using the diagonal part of the Hamiltonian truncated to terms $J(g)_{diag}^{(k-1)}$ of order less than L_{k-1} (see p. 48 of paper¹). Such terms are diagonal, thus they have to wrap onto itself at least once to cancel all the flips. Using that, Imbrie proves a strong bound on extra support on each side $c_k = \frac{1}{2} \text{supp} E_{\sigma}^{(k-1)} - E_{\sigma'}^{(k-1)} \leq \frac{15}{14} L_{k-1}$. So the term we get in $A(g)$ has the same positions of spin flips as $J(g)$, and a collar around those spin flips of size $\frac{15}{14} L_{k-1}$.

Let's now discuss the support of $J^{(k-1)}(g)$ included in $A^{(k)}(g)$. The orders of terms $A^{(k)}(g)$ are in $[L_k - 1, L_k)$. The order $|g|$ limits the size the interval containing its spin flips. Thanks to bridging of the gap described on pp. 45-46 of paper¹, order directly translates into maximal size of the interval where flips are happening (it is exactly $|g| < L_k$ for terms in $A^{(k)}$, and $|g|/2 \leq L_{k-1}/2$ for diagonal terms in $J(g)_{diag}^{(k-1)}$ appearing in the denominator for $A^{(k)}$). So the total support is either $|g|$ or $|g|/2$ plus two collars c_k . We illustrate how the collars appear by explicitly listing the properties of terms for the first few steps.

$L_0 = 1$	perturbation order $[1, 15/8]$ is removed at step 1 by $A^1 \Rightarrow$	$H^1 + J^1$
$L_1 = 15/8$	perturbation order $[2, 3]$ is removed at step 2 by $A^2 \Rightarrow$	J^1 starts with order 2
	H^2 includes order 2, support 3 terms	$H^2 + J^2$
$L_2 \approx 3.5$	perturbation order $[4, 6]$ is removed at step 3 by $A_{c=2}^3 \Rightarrow$	J^2 starts with order 2
	H^3 includes order 6, support 7 terms	$H^3 + J^3$
$L_3 \approx 6.6$	perturbation order $[7, 12]$ is removed at step 4 by $A_{c=5}^4 \Rightarrow$	J^3 starts with order 7
	H^4 includes order 12, support 16 terms	$H^4 + J^4$
$L_4 \approx 12.4$	$c = 6 + 3 + 1 + 1 = 11$	J^4 starts with order 13

We see that even though in H^3 the support 7 terms are present, somehow the collar for A^4 is chosen to be 5. In other words, the path in the collar at every step should be touching, otherwise the collars blow up. We assume that terms with bigger collar are not included in each $A(g)^{(k)}$. Let us see what is the order of the extra terms that would appear if we tweak the procedure this way — if it is the same order as $J^{(k)}$, then we can just include them with other paths.

For illustration, let's use denominators within a shorter collar in A^4 , and check that $e^A(H^3 + J^3)e^{-A} = H^4 + J^4$ with the same condition on J^4 to include only $|g| > 13$. Recall that we want to remove terms in J^3 of orders 7 to 12. Now in principle H^3 contains order 6, support 7 terms, that look like 2-site collar around the graph of step 3. The collar itself is a step 2 graph of support 3, order 2, with collar of size 1 around it. It doesn't go further down. Now after conjugation, $[A, H^3]$ should cancel terms in J^3 in such a way that only terms of order > 12 remain. But instead of cancellation, it gives a $(1 - \Delta E_{full}/\Delta E_{short})J_{7..12}^3$ remainder $\Delta E_{full} - \Delta E_{short}$ contains terms in H^3 whose center graph does not touch the spins flipped by A . In this case, it is only the center size 3, collar 2 (so support 7) term of order 6. $7 + 6 = 13$ which is enough to send it to J^4 , and it's not even enough to make it connected - one needs to lift an order 2 term to make it continuous. Unfortunately this activity does not really fit into the structure that Imbrie proves bounds for, but it can be done at least in principle. Since Imbrie makes no mention of this, we leave the exact way it's done as an open question and proceed.

Here's the table at k 'th step:

$$\begin{array}{ll}
L_{k-1} = (15/8)^{k-1} \text{ pert. order } [\text{ceil}(L_{k-1}), \text{floor}(L_k)] \text{ removed at step } k \text{ by } A_{c_k}^k \Rightarrow & H^k + J^k \\
L_k \approx (15/8)^k & \begin{array}{l} H^k \text{ includes order } \text{floor} L_k, \text{ support } \text{floor}(L_k/2) + 2c_k \text{ terms} \\ J^k \text{ starts with order } \text{floor} L_k + 1 \\ c_{k+1} = \text{floor}(L_k/2) + c_k \end{array}
\end{array}$$

In other words $c_{k+1} = 1 + \sum_{k'=2}^k \text{floor}(L_{k'}/2) \leq \sum_{k'=1}^k L_{k'}/2$ which is the expression used by Imbrie. Let's evaluate it:

$$c_{k+1} = \frac{1}{2} \sum_{k'=1}^k (15/8)^{k'} = \frac{1}{2} (15/8)^k \sum_{k'=0}^{k-1} (8/15)^{k'} \leq \frac{1}{2} L_k \sum_{k'=0}^{\infty} (8/15)^{k'} = \frac{1}{2} L_k \frac{1}{1 - 8/15} = \frac{15}{14} L_k \quad (\text{B35})$$

This is exactly as Eq 4.50 on p.46. So the support of a term in A^k is bounded by $\text{floor} L_k + 2c_k \leq L_k + \frac{15}{7} L_{k-1} = \frac{15}{7} L_k = \frac{8}{7} L_{k+1}$. And the order is bounded by L_{k-1} from below as seen in Subsection B1 (even $k = 1$ works out since $\text{ceil} L_0 = 1$). The final piece we need for the postulates in the main text is the support of resonances: after perturbative rotations A^k of step k , there come nonperturbative rotations that may flip spins within an interval $3L_k$ (p.15) and have a collar $15/14L_j = 15/14L_{k-1}$ (p.25) on each side. So total support is $3L_k + \frac{8}{7} L_k = 4\frac{1}{7} L_k \leq 4.2L_k$.